International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)
ISSN(P): 2249-6890; ISSN(E): 2249-8001
Vol. 8, Issue 1 Feb 2018, 1135-1150
© TJPRC Pvt. Ltd.

EFFECTIVE TIME DOMAIN FEATURES FOR IDENTIFICATION OF BEARING FAULT USING LDA AND NB CLASSIFIERS

B R NAYANA[1] & P GEETHANJALI[2]

[1] *Department of Electrical & Electronics Engineering, Sir MVIT, Bangalore, India*

[2]*School of Electrical Engineering, VIT University, Vellore, India*

ABSTRACT

Recently, the mechanical fault detection of an induction motor (IM) from vibration signals using pattern recognition has proven to be an effective method. This paper has studied for the first time statistical time domain features mean absolute value (MAV), waveform length (WL), zero crossing(ZC), slope sign changes (SSC), simple sign integral(SSI) and Willison amplitude (WAMP) for identification of the mechanical faults using linear discriminant analysis (LDA) and naive Bayes (NB) classifiers. In this study, the effectiveness of the features is investigated using parameters like accuracy, sensitivity and specificity individually and in groups for a total of 63 combinations. Each feature set combination is investigated for 15datasets defined under 5 groups in different combinations of faulty and normal working conditions. The results indicate that the feature set of SSI, WL,SSC and ZC features outperform the conventional features in the identification of faults and is found to be computationally effective. Further, NB classifier is found to be better than LDA in identification of mechanical faults.

KEYWORDS: Fault Diagnosis, Statistical Features, Linear Discriminant Analysis Classifier, Naive Bayes Classifier, Roller Bearings & Pattern Recognition

Original Article

Received: Jan 05, 2018; **Accepted**: Jan 25, 2018; **Published**: Feb 07, 2018; **Paper Id**.: IJMPERDFEB2018135

1. INTRODUCTION

In recent years, there has been considerable evolution in the field of fault diagnosis of induction machines with the aid of expert systems and artificial intelligence algorithms. Many condition monitoring techniques have been successfully developed and implemented. Bearing faults are among the more prominently occurring faults [1] and hence their diagnosis forms an essential part in condition monitoring of induction machines. Large number of detection techniques have been developed based on signature analysis of either stator current or vibration signals. Among this, vibration signals have been proven to be more reliable for diagnosing mechanical faults either invasively or non-invasively. Condition monitoring of bearing faults with pattern recognition involves feature extraction, feature selection, feature reduction and their classification. Typically, pattern recognition methods are applied to diagnose the faults with time domain features like peak value, crest factor, kurtosis, etc.[2-3].Prior researches in this area using time domain features like mean, standard deviation, shape factor, etc. have been found to yield poor results [4]. Investigations using frequency domain features like power spectrum, power spectral density, periodograms etc.[5-6] relies on the differences in frequency characteristics of fault conditions[7]. These differences are non-significant and hence difficult to diagnose. As vibration signals are non-stationary in nature, time–frequency domain analysis like spectrogram, wavelets transforms(WT) etc. have been used for extracting features to identify the bearing faults[7-12]. This analysis using WT methodology [13] suffers a major setback due

to adjustable windowed Fourier transforms energy leakage occurring during signal processing. Another limitation of this technique is that the success of this relies heavily on the choice of appropriate base function which determines the frequency bands of the decomposed signals.

In the present work, the authors attempted on novel time domain features such as mean absolute value (MAV), waveform length (WL), zero crossing (ZC), slope sign changes (SSC), simple sign integral(SSI) and Wilson amplitude(WAMP) and found time domain features outperform frequency and time-frequency features. Though, frequency and time-frequency features necessitate the dimensionality reduction methods prior to classifiers. These time domain features do not require any feature reduction or selection schemes and hence found to be computationally effective. Literature survey shows that various classification techniques such as k-nearest neighbor (KNN) [14,15] artificial neural network (ANN)[3,16,17] support vector machine[18], linear discriminant analysis (LDA)[2] etc. have been employed to study the performance of the extracted features. In this paper, the authors have used naive Bayes (NB) classifier and compared the results thus obtained for variation in parameters like accuracy, sensitivity and specificity with that obtained using LDA classifier. The results evince that time domain features identify the bearing faults with good accuracy compared to other features considered in the literature [2], [13-14], [17-22]. Overall 63 feature set combinations from 6 features have been employed for bearing fault diagnosis of 5 groups of data involving 15 datasets which has been drawn in combinations of location of fault and load condition. Though condition monitoring schemes of bearing faults involve Feature extraction, feature selection, features reduction and classification processes which will be handled by various methodologies like WT for feature extraction and minimum-redundancy maximum-relevancy method for feature selection and differential evolution algorithm for classification[21] and spectrum imaging has been implemented for feature extraction and enhancement[22] in previous works. The present work focuses on the simplest scheme development to serve the purpose; hence Feature extraction and classification alone are implemented. It should be noted that the approach is not limited to these 6 features. Other features, such as MAV slope, log detection, peak factor, etc. can be also used. Detailed discussions of which type of features are more useful than others for bearing fault diagnosis are beyond the scope of this paper.

2. METHODOLOGY

2.1. Dataset Description

The vibration recordings from experiments conducted using a 2 HP Reliance Electric motor by CWRU (Bearing vibration dataset, 0000) has been used to derive five groups (A-E). The drive end (DE) bearing (6205-2RS JEM SKF make) and fan end (FE) bearing (6203-2RS JEM SKF bearing) are selected as the test bearings. Motor bearings were seeded with faults using electric discharge machining (EDM). Fault depths of 0.007 inch, 0.014 inch and 0.021inch with 0.040 inch of diameter were artificially created at the inner raceway (IR), rolling element (RE) and outer raceway (OR). Faulted bearings with respect to all 3 faults (3F) were reinstalled into the test motor and vibration data was recorded for motor loads of 0 to 3 HP (motor speeds of 1797 to 1730 RPM) individualistically. Vibration data have been collected using accelerometers, which were placed at 12 O' clock position at DE and FE of the motor housing at a sampling rate of 12kHzfor a duration of 10 seconds. The data recorder had been equipped with a low - pass filter at the input stage for antialiasing. The benchmark study made by [23] indicates that the central load zone is at 6 O' clock position. In addition, the study revealed the fact that some of the vibration signals recorded on DE and FE positions were non useable due to lack of clarity, clipping of the data, and contamination of data due to the presence of significant electrical noise signals. The five groups (A-E) derived are as shown in Table1 with varying combinations of bearing fault depth in milli inches (FD)

and load conditions to explore the effectiveness of time domain features and classifiers considered in this study.

Overall 8 normal and 60 faulty working conditions are used for the analysis. The faulty conditions considered are obtained from DE data for 3F with 3 FDs and FE data for 3F with 2 FDs, each under 4 different loading conditions. Correspondingly, 8 normal working conditions for DE and FE are considered each of 4 different loading conditions. In Group A, the datasets A-i, iii and v are derived to study a four class classification of N and defects with IR, RE and OR for identical load and identical FD conditions. Dataset A-i is drawn for DE N and all 3F with FD of 7, thus includes 4 working conditions for each load respectively.

Table 1: Basic Information of 5 Groups

Group		Dataset Description	No of Working Conditions	No of Classes
A	i.	DE-N and 3F of FD 7 for each load	4	4
	ii.	DE-N and 3F of FD 7 for all 4 loads together	16	
	iii.	DE-N and 3F of FD 14 for each load	4	
	iv.	DE-N and 3F of FD 14 for all 4 loads together	16	
	v.	DE-N and 3F of FD 21 for each load	4	
	vi.	DE-N and 3F of FD 21 for all 4 loads together	16	
B	i.	DE-N and 3F of 3FDs (7,14, 21) for all 4 loads together	40	4
	ii.	DE- N and 3F of 3 FDs (7,14, 21) for each load	10	
	iii.	DE- N and 3F of 2 FDs(7 & 21) for all 4 loads together	28	
	iv.	DE- N and 3F of 2 FDs (7 & 21) for each load	7	
C	i.	DE-3F of 2 FDs(7 & 21) for each load	6	6
	ii.	FE-3F of 2 FDs(7 & 21) for each load		
D	iii.	DE & FE-3F of FD 7 for each load	6	6
	iv.	DE & FE 3F of FD 21for each load		
E		DE- N and 3F of 3 FDs (7,14, 21) for each load	10	10

Datasets A-iii and A-v are derived in a similar manner for FDs of 14 and 21 respectively. A four class classification is studied in datasets A-ii, A-iv and A-vi for identical FD conditions, but being independent of load with FDs of 7,14 and 21 respectively. Thus, datasets A-ii, A-iv and A-vi includes 16 working conditions. Group B, is also derived for the same four class classification, with different combination of FDs and load conditions. The datasets B-i and B-ii deals with FDs of7, 14 and 21; and FDs of 7 and 21 are considered for datasets B-iii and B-iv. However, dataset B-i and B-iii are implemented irrespective of load condition, thus 40 and 28 working conditions are employed correspondingly. Whereas, datasets B-ii and B-iv are implemented with 10 and 7 working conditions respectively for each load condition. Group C considers the working conditions of same bearings with all 3Fof FD7 and 21to perform a 6 class classification for every load condition. Hence, group C includes 2 datasets, one for DE and one for FE respectively. A similar analysis is performed with group D wherein, the datasets are for all 3Fwithsame FD over DE and FE. Therefore one dataset is for FD of 7 and other for FD of 21 respectively. Group E, includes N and all 3F for FDs of 7,14 and21 of DE for every load condition. Therefore, 10 working conditions are employed of for a 10 class classification.

2.2. Feature Extraction

The temporal characteristics, hidden in the vibration signals are extracted using newly attempted time domain features such as mean absolute value, simple sign integral, waveform length, Willison amplitude, zero crossing, and slope sign for bearing fault diagnosis. The mathematical description of proposed features is presented in this section.

Mean Absolute Value (MAV): Mean absolute value is the average of absolute value of data for a segment of

length L and is defined in equation (1). MAV is similar to average rectified value and can be calculated using the moving average of full-wave rectified vibration signal.

$$MAV = \frac{1}{L}\sum_{i=1}^{L} |y[i]| \tag{1}$$

Simple Sign Integral (SSI): Simple sign integral is the integral of square of data samples. It determines the energy of the data segment and is computed using equation (2).

$$SSI = \sum_{i=1}^{L} |y[i]|^2 \tag{2}$$

Waveform Length (WL): Waveform length is the cumulative length of the waveform over the time segment. It is related to amplitude, time and frequency information of the data segment and is calculated using equation (3).

$$WL = \sum_{i=1}^{L} |y[i] - y[i-1]| \tag{3}$$

Willison Amplitude (WAMP): Willison amplitude is the number of times the difference between amplitude of adjacent samples exceeds a predefined threshold value. It is calculated using equation (4) and (5).

$$WAMP = \sum_{i=1}^{L} \phi(|y[i] - y[i+1]|) \tag{4}$$

$$\text{where } \phi(x) = \begin{cases} 1 & if\ x \geq \in \\ 0 & otherwise \end{cases} \tag{5}$$

and $\in$ is the threshold value and chosen as 0.5.

Zero Crossing (ZC): Zero crossing is the number of times the signal crosses zero. This is a feature, provides information about frequency of the signal and is calculated from (6) which satisfy equation (7).

$$ZC = \begin{cases} (y[i] > 0 \&\& y[i+1] < 0) \\ or \\ (y[i] < 0 \&\& y[i+1] > 0) \end{cases} \tag{6}$$

$$|y[i] - y[i+1]| \geq \epsilon \tag{7}$$

To abstain from the background noise a small threshold of ϵ =*0.5* is chosen.

Slope Sign Change (SSC): Slope sign change is another feature that characterizes the frequency and is computed using equation (8) and satisfying equation (9).

$$SSC = \begin{cases} (y[i] > y[i-1] \&\& y[i] > y[i+1]) \\ or \\ (y[i] < y[i-1] \&\& y[i] < y[i+1]) \end{cases} \tag{8}$$

$$|y[i] - y[i-1]| \geq \epsilon \tag{9}$$

Slope sign change indicates the number of changes between positive and negative slope among three consecutive segments. A threshold ϵ =0.5 is used for avoiding the interference in vibration signal.

These 6 features extracted from the bearing vibration signals are given as input for LDA and NB classifiers. The effectiveness of the features is studied in 63 feature set (FS) combinations.

$$FS=\{FS_1,FS_2,FS_3,FS_4,FS_5,FS_6,FS_7,FS_8\ldots\ldots\ldots\ldots..,FS_{61},FS_{62},FS_{63}\} \tag{10}$$

where FS_1-FS_6 are FSs of individual features, FS_7 is set of MAV,SSI and FS_8 is set of MAV,WL similarly further FSs are derived with combinations of 2,3,4,5 and 6 features. The performance of these time domain features has been investigated using vibration data recorded by Case Western Reserve University (CWRU) [24]. Each working condition of the dataset contains 10 seconds of vibration signal form DE and FE from which, 65536 samples (5.46 seconds) are considered for processing and are segmented into windows of length 1024 with 50% overlapping. Features are extracted for each segment thus resulting in a feature length of 128 for every feature pertaining to respective working condition. Based on the experimental motor bearing data discussed in section 2.1, the analysis results are drawn in section 3 to verify the effectiveness of these 6 features in 63 combinations for bearing fault diagnosis, with respect to accuracy, sensitivity and specificity.

2.1 Classification

LDA classifier

Linear discriminant analysis is the most common technique used for data classification and dimensionality reduction. Linear discriminant analysis easily handles the case where the within-class frequencies are unequal and their performances have been examined on randomly generated test data. LDA approach for classification considers posterior probability, prior probability and cost of classifying an observation to a particular class. Thus the objective is to minimize the classification cost and the minimization function is defined as

$$\hat{c} = arg \min_{c=1\ldots N} \sum_{k=1}^{N} \hat{P}(k|X) cost(c|k) \tag{11}$$

where

$\hat{c}$ is the predicted class.

N is the number of classes.

$\hat{P}(k|X)$ is the posterior probability of class k for observation X.

cost(c|k) is the cost of classifying an observation as c when its true class is k.

The posterior probability that an observation X belongs to class k given as

$$\hat{P}(k|X) = \frac{P(X|k)P(k)}{P(X)} \tag{12}$$

Where P(k) represents the prior probability of class k.

P(X) is a normalization constant, that is, the sum over k of P(X|k)P(k)

P(X|k) is the multivariate normal density function and is given as

$$P(X|k) = \frac{1}{(2\pi|CM_k|)^{1/2}} \exp\left(-\frac{1}{2}(X - \mu_k)^T {CM_k}^{-1}(X - \mu_k)\right) \tag{13}$$

Where $CM_k$is the covariance matrix of k^{th} class.

$$CM_k = (X_{k-}\mu_k)(X_{k-}\mu_k)^{\ T} \tag{14}$$

and μ_k is the mean of k^{th} class.

The LDA classifier steps can be summarized as to estimate the prior probabilities, mean and covariance matrix for each class. Further, for a new observation X estimate the class using equation (11).

NB classifier

Naive Bayes is based on Bayes theorem suited to solve the high dimensional problems. Parameter estimation for naive Bayes models uses the method of maximum likelihood and performs better in many complex real world situations

The advantage of NB classifier is it requires a small amount of training data to estimate the parameters. The algorithm for implementation of NB classifier is as follows:

If there are 'm' classes: $C_1, C_2, C_3 \ldots C_m$, and the feature vector X : $[x_1, x_2, x_3, \ldots x_n]$, for n number of features, the naive Assumption of class conditional independence computed using equations (15) and (16).

$$P(X/C_i) = \prod_{k=1}^{n} P(x_k/C_i) \tag{15}$$

$$P(X/C_i) = P(x_1/C_i) * P(x_2/C_i) * P(x_3/C_i) * P(x_4/C_i) * .. P(x_n/C_i) \tag{16}$$

NB classifier predicts that X belongs to Class Ci iff

$$P(C_i/X) > P(C_j/X) \text{ for } 1 <= j <= m, j <> i \tag{17}$$

The maximum posteriori hypothesis can be stated as

$$P(C_i/X) = P(X/C_i)\, P(C_i)/P(X) \tag{18}$$

Maximize $P(X/C_i)P(C_i)$ as P(X) is constant. (19)

where $P(C_i)$ is class prior probability.

P(X) is the prior probability of X.

$P(C_i/X)$ is the posterior probability.

$P(X/C_i)$ is the posterior probability of X conditioned on Ci.

With many attributes, it is computationally expensive to evaluate $P(X/C_i)$. Being conditionally independent and computationally expensive are the only drawbacks of this classifier.

2.2 Performance Metrics

Classification of bearing conditions for groups A to E are employed with 63 feature set combinations and the performance is assessed for 50% training and 50% testing sizes respectively. The performance is evaluated for all FS combinations based on accuracy, sensitivity, specificity of LDA and NB classifiers. Sensitivity, SE is defined as the rate of overall number of true positives (TP) (correctly classified patterns) to the total number of actual positive patterns (TAp)

$$\text{Sensitivity} = {}^{TP}/{}_{TA_p} \tag{20}$$

Specificity, SP is defined as the rate of total number of true negative (TN) to the total number of actual negative patterns(TAn) :

$$\text{Specificity} = {}^{TN}/{}_{TA_n} \tag{21}$$

The overall accuracy AC, is estimated as the percentage of rate of TP and TN to total number of patterns, N under consideration for classification.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{N}} * 100 \quad (22)$$

However, the overall accuracy contributed by every feature depends even on positive prediction value and negative prediction value, that is, even if sensitivity is 1 or specificity is 1 accuracy is not necessary to be achieved as 100%.

3. RESULTS

The classification performance of a classifier is investigated with 3 parameters in this work as discussed in section 2.4. Hence,Figure.1 and figure.2 shows a plot of Accuracy, Sensitivity and specificity as a function of the FS number for a four class classification of N, IR, OR, RE with a fault depth of 7,14 and21 using LDA and NB classifiers respectively for group A. It is observed that the accuracy is 100% of the data sets I and II for all individual features except ZC and SSC. The performance of these features does not improve even in combined form. In data set iii the accuracy is 40% to 80% with the same features which had excelled in performance for data set I and ii whereas the features ZC and SSC have given accuracy of 95-100%. However in data set v performances of all features are better than dataset iii, and it is noticed that the feature WL, which had given the least accuracy in dataset iii exhibits maximum accuracy in this dataset. It is seen that as the number of features grouped increases the performance also improves and the accuracy reaches 100%. It can be witnessed that ZC and SSC are the features resulting in maximum accuracy when implemented in combination with any other feature for datasets iii to vi. Table 2, presents the number of features required for attaining maximum accuracy for every dataset of group A. It is interesting to note that both in LDA and NB classifications as seen in figure 1 and 2, the sensitivity and specificity dips for the features ZC and SSC, both for individual and combined cases thus resulting in poor accuracy levels in the classification. But ZC and SSC when combined with any other feature will excel in performance and exhibit maximum accuracy. Therefore, it is important to analyze all parameters of a feature before either considering or rejecting it for classification. The classification with NB as shown in figure 2,shows better results when compared to LDA in figure 1, as the patterns do not get scattered.

Table 2: Number of Features Required to Attain Maximum Efficiency for the Datasets of Group A.

Group	A														
Dataset	DE-N&3F-7D					DE-N&3F-14D					DE-N&3F-21D				
Load	4L	L-0	L-1	L-2	L-3	4L	L-0	L-1	L-2	L-3	4L	L-0	L-1	L-2	L-3
Max AC	100	100	100	100	100	100	100	100	100	98	100	100	100	100	100
No of Features	1	1	1	1	1	3	3	1	2	3	3	2	2	1	1

Table 3: Number of Features Required to Attain Maximum Efficiency for the Datasets of Group B

Group	B									
Dataset	DE-N, 3F, 3 FD(7,14,21)					DE-N, 3F, 2 FD(7,21)				
Load	4L	L-0	L-1	L-2	L-3	4L	L-0	L-1	L-2	L-3
Max AC	91.4	97.2	98.3	90	92.3	100	100	100	100	100
No of Features	4	5	4	5	3	3	2	2	4	3

Figure 3, (i) presents the accuracy of classification for 4 datasets using LDA and NB. It is observed that for LDA though many FSs exhibit 100% accuracy the average AC of NB is greater than LDA. The reason is, when compared to NB

sensitivity of LDA is less and specificity remains to be almost same with respect to each FS, thus the numbers of false positive patterns are more with LDA classifier. Further, it is observed that datasets B-iii and B-iv exhibits 100% and requires less number of features when compared to datasets B-i and B-ii as shown in table 3. This is for the known fact of resonant vibration signals interfering with the vibration signals of FD 14, as discussed earlier for the observations made by group A.

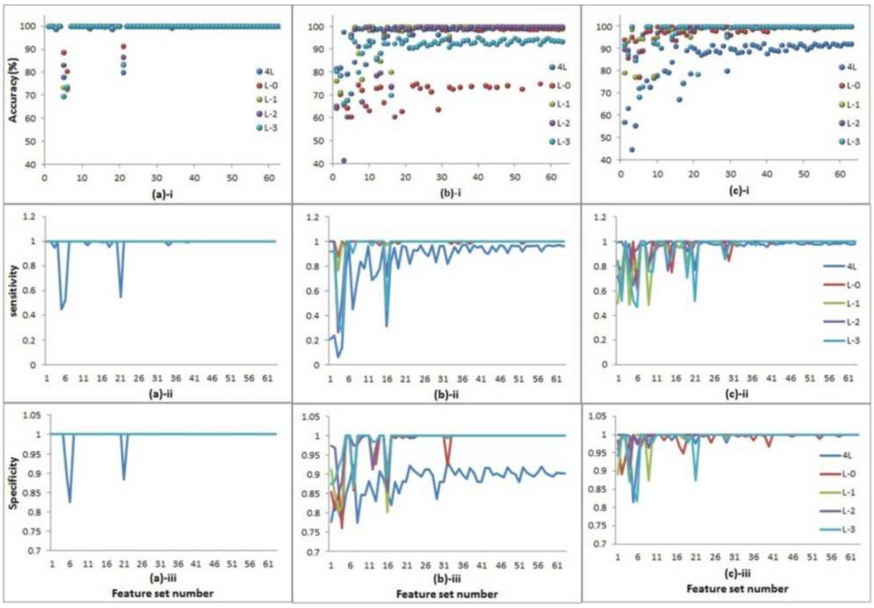

Figure 1: Accuracy, Sensitivity and Specificity as a Function of Feature Set Number for a Four Class Classification of N, IR, OR and RE with (a) FD-7, (b) FD-14 and (c) FD-21 Using LDA for Group A.

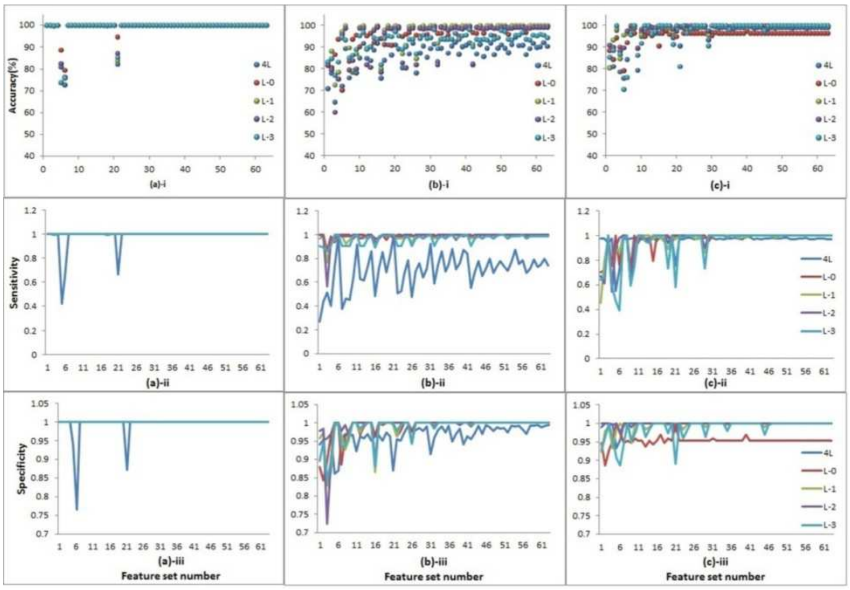

Figure 2: Accuracy, Sensitivity and Specificity as a Function of Feature Set Number for a Four Class Classification of N, IRD, ORD, BD with (a) FD-7, (b) FD-14 and (c) FD-21 Using NB for Group A.
In General, NB Is Better than LDA.

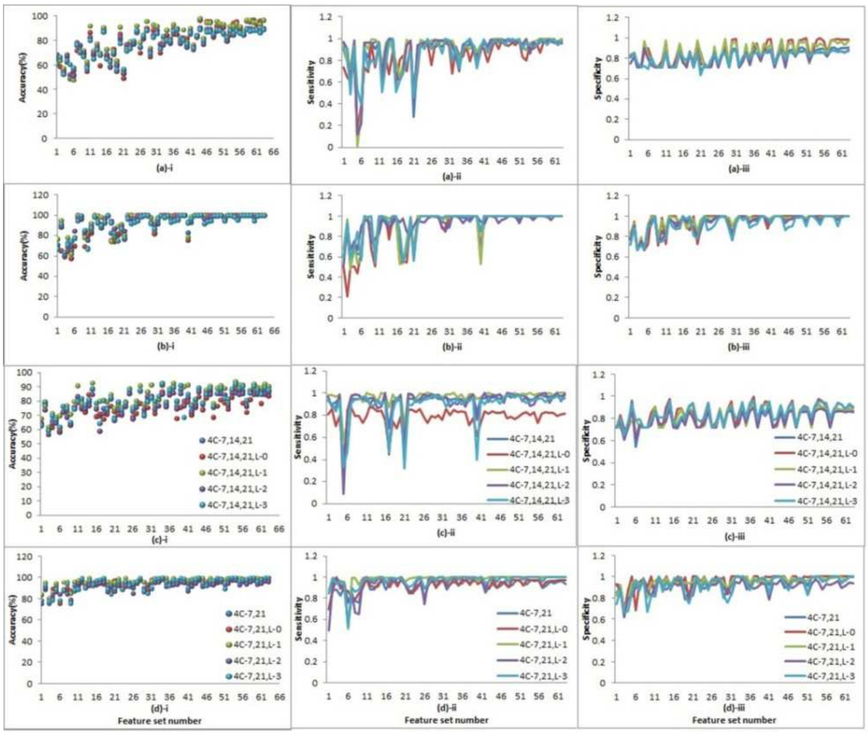

Figure 3: (i)Accuracy, (ii)Sensitivity and (Iii)Specificity as a Function of Feature Set Number for a Four Class Classification of N, IR, OR, RE is Dataseted Irrespective of Load, and with Respect to Each Load for (a) FD-7,FD-14 and FD-21, and (b) FD-7 and FD-21 Using LDA, (c) and (d) Using NB for Group B. In General, NB is Better than LDA.

In figure 3, (a) and (b) are for LDA and; (c) and (d) are for NB, which clearly shows LDA has scattered accuracy patterns along with less sensitivity, whereas average specificity remains same for LDA and NB respectively. Figure 4, shows accuracy, sensitivity and specificity for group C and it is observed that LDA performance is slightly better than NB. In addition to this, it can be observed that for NB sensitivity does not show considerable improvement as number of features combined increases like in other groups. However, its value is less in FSs where feature WL, WAMP, ZC exist individually or in combination. Mainly when feature WL, WAMP and WL, ZC combination exists with any other feature poor sensitivity and is attained. Conversely, specificity is good for all FSs when compared to LDA, as it improves with the number of features combined and reaches maximum at the earliest. Therefore NB classifier performs better than LDA except for load 0HP condition. Table 4 presents the number of features required to attain maximum accuracy for datasets of group C. It is clear from the table that good accuracy ranges are obtained for dataset C-i, as it deals with DE data. However, in datasetC-ii the accuracy ranges have not reached 100% with any FS, as it deals with FE data. In which, certain data files are either non diagnosable or having electrical noises or the signal is clipped off as stated in the benchmark study made by [23]. Therefore, it can be observed from figure 4 that the accuracy of classification is less for dataset C-ii of load 2HP condition and the same can be seen in figure 5, for dataset D-i of load 2HP condition. As C-ii and D-i of load 2HP condition employs the same file which has more noise, the accuracy of classification is less. However, when the signals were processed for complete 10 seconds instead of 5.4 seconds as considered in this work earlier, the accuracy levels are considerably improved as seen in table 4 and 5 for the respective datasets. Table 5 presents the number of features required to attain maximum accuracy for datasets of groups D and E. For the datasets of group E accuracy, sensitivity and specificity are plotted figure 6, and the figure illustrates that among the existing 63 FS, the dips are formed due to the FSs

WL and WAMP. It can also be observed that though sensitivity and specificity for the feature SSC is 1 for all load conditions with both classifiers the accuracy is not 100%. For the reason, that positive prediction and negative prediction values are not sufficiently enough. Overall, for all datasets the FS of SSI, WL, SSC and ZC can give the maximum accuracy. However, the FS of WL, SSC and ZC features is sufficient for most of the dataset to attain maximum accuracy and for some datasets single feature is sufficient as seen in figure 7. Therefore the present approach is simple and computationally cost effective.

Table 4: Number of Features Required to Attain Maximum Efficiency for the Datasets of Group C

Group	C							
Dataset	i-DE-3F-7D&21D				ii-FE-3F-7D&21D			
Load	L-0	L-1	L-2	L-3	L-0	L-1	L-2	L-3
Max AC	97	100	100	100	95.05	94.01	90.1	97.4
No of Features	3	3	2	2	4	2	4	4

Table 5: Number of Features Required to attain Maximum Efficiency for the Datasets of Group D & E

Group	D								E			
Dataset	i-DE&FE-3F-7D				ii-DE&FE-3F-21D				DE-N, 3F, 3 FD(7,14,21)			
Load	L-0	L-1	L-2	L-3	L-0	L-1	L-2	L-3	L-0	L-1	L-2	L-3
Max AC	99.5	97.14	93	100	99.74	98.7	98	99.5	97.3	91.95	91.83	98.3
No of Features	3	2	4	3	3	4	3	3	3	4	4	4

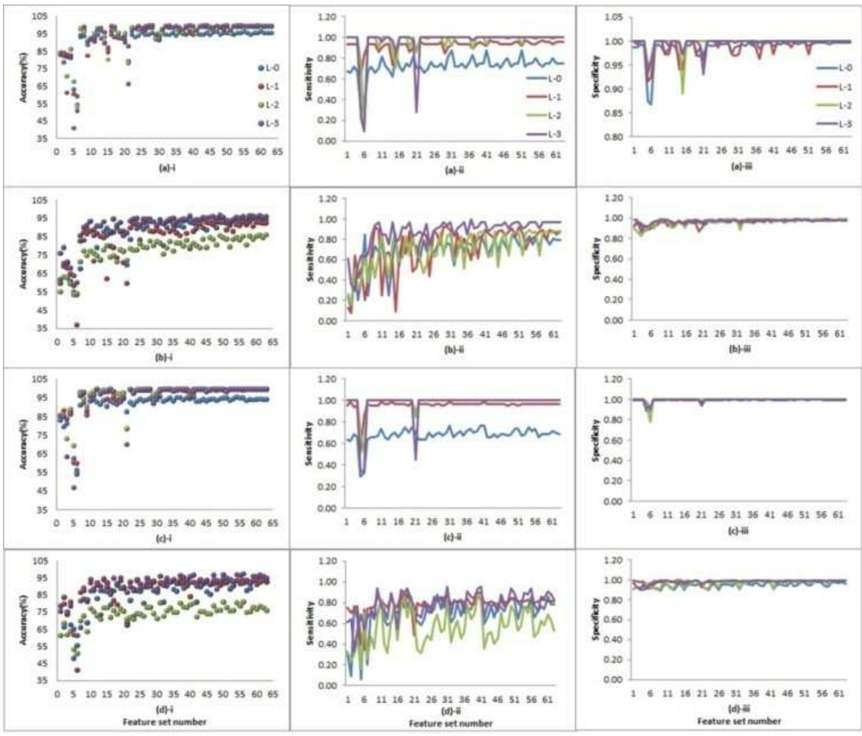

Figure 4: Average (i)Accuracy,(ii)Sensitivity, (iii)Specificity of Group C, for a 6 Class Classification of all 3 Faults for DE with FD of 7 and 21 (a) and for FE (b) Using LDA, Similarly (c) and (d) Using NB. In General, NB Is Better than LDA

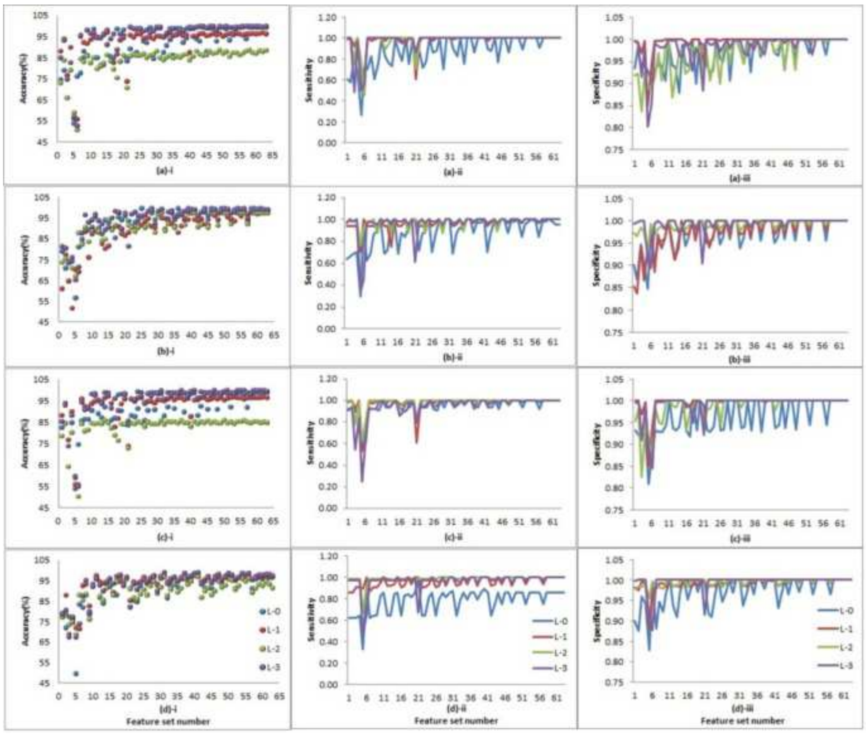

Figure 5: Average Accuracy, Sensitivity and Specificity of Group D, for a 6 Class Classification of all 3 Faults for DE and FE with FD of 7 (a) and 21(b) Using LDA, Correspondingly (c) and (d) Using NB. Overall, LDA Performed Better than NB

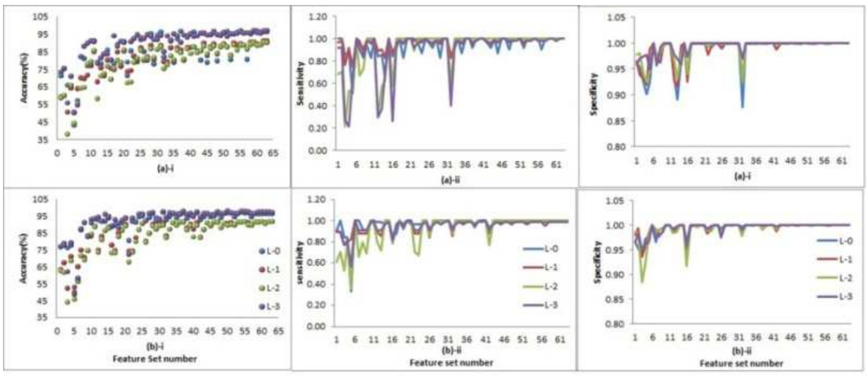

Figure 6: Accuracy of a 10 Class Classification, of DE with all 3 Faults and all 3 FDs Including Normal Working Condition of Group E, Using LDA and NB

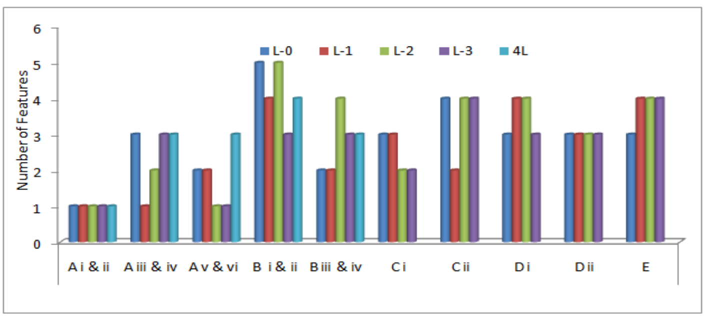

Figure 7: Number of Features Required for each Dataset to Attain Maximum Accuracy with Respect to Load Conditions

4. DISCUSSIONS

The statistical time domain features MAV, SSI, WL, WAMP, ZC, SSC is introduced for bearing fault diagnosis using LDA and NB classifiers. Even though some studies have been performed [3], where time domain features are applied but the feasibility of above discussed features for bearing fault diagnosis has not been investigated so far. Our findings with the datasets derived with a new set of features are in agreement with the previous studies. The investigations performed disclose the fact that features considered will perform well for the combination of features associated with time and frequency. To be precise, the feature SSC is associated with time and the feature ZC is associated with the frequency of a signal, when these 2 features are employed together then it leads to maximum accuracy. However, the parameter sensitivity, specificity, positive prediction and negative prediction values and others to be considered before choosing the combination of features to develop the scheme for automated bearing fault diagnosis, which can provide best discrimination with less computation time. It is perceived that features ZC and SSC exist as main features in determining the maximum accuracy for almost all the data sets discussed above. Either feature ZC and SSC together or ZC and SSC along with WL will give maximum accuracy. But, overall4 features are sufficient for the authors to get maximum accuracy in all datasets of groups A to E. Though in dataset B for certain load conditions 5 features are providing maximum accuracy, it is not exhibiting considerable improvement. Therefore the FS of SSI, WL, SSC and ZC features will provide maximum accuracy for all datasets of group A-E. The application of these features for bearing fault diagnosis has a number of benefits over other methods proposed so far. Such as, the features selected are one dimensional simple and fast to estimate and also they are less in number. This avoids the need for feature selection and reduction processes. Secondly, classifications can be implemented by simple classifiers like LDA and NB classifiers; hence diagnosis can be implemented with fewer computations. Altogether these factors constitutes that the proposed method is highly appropriate for real-time analysis. Another important advantage commonly believe is, these features provide same information as time, frequency and time-frequency analysis of the signals as implemented by[2-4][8-9].

For the comparison between results obtained from the proposed method and the existing methods in literature, only the works which have used identical dataset are considered. However, the differences exist with respect to the vibration data considered is 12kHz or 48kHz and the number of channel inputs taken into consideration. Few authors have chosen single channel input, either DE or FE data like in present work for diagnosis and some authors consider both DE and FE data for every working condition. The accuracy obtained from the proposed method gives the best accuracy for group A, and it is equivalent to the best presented. This is in corroboration with earlier reporting's[18, 21]. Specifically datasets A-II, which is a four class classification for FD of 7 being, irrespective of load is implemented by authors in [21] excluding load 0HP condition. Further WPD for feature extraction and mRMR for feature selection and DE-EAM for classification are employed and average accuracy of 96.1% is obtained. Similarly spectrum imaging and feature enhancement is applied for feature extraction, and classification is realized using ANN by the authors in [22] for dataset A-i with 2HP load and obtains an accuracy of 96.9%. Data setA-iii is assessed for load 0, 1 and 2HP using 20 features by authors in [18] and attained an accuracy of 99.56%,100%,99.89% respectively. But100% accuracy is obtained in the present work for all datasets of group A by using a maximum of 3 features. For group B, identical implementation of dataset B-i is performed by authors in [17] and in [19]. In the former work the authors would generate feature vectors using characteristics of the feature ZC and classify the fault using Feed forward NN and results indicates for 20 ZC intervals accuracy obtained is 92.45% using a window length of 1024. In the later work, the authors have made use of NPE, SOM along with many classifiers including LDA. Further, they have extracted10 time domain and 10 frequency domain features

to realize using classifiers and have attained an accuracy of 96% without denoising and 98% after denoising the data for LDA classifier respectively. For B-iii dataset authors of [20] have implemented by fuzzy inference methodology and by excluding normal working condition besides obtains a maximum accuracy of 73%. But in the present work, 100% accuracy is obtained either when implemented irrespective of load or with respect to each load, as shown in table 2 for B-iii and B-iv datasets. However, though accuracy does not excel for all cases, they are in substantiation with earlier works.

The comparison of datasets C and D group of present work and identical implementation in[13]can be discussed for every load condition. The authors have extracted 5 IMFs by EEMD and classified with SVM classifier It is seen from table 6that the present work has performed better for datasets C-i and D-ii. But for datasets of C-ii and D-i are in par with earlier work with a difference of 0% to 2% except for the cases in which electrical noise is assumed to be present as discussed earlier. However, in the present work the window length is 1024 and 5.4 seconds of data are processed for every working condition, whereas in [13] the window length is 3000 and 10seconds of complete data is processed.

Table 6: Comparison of Present Work and X. Zhang et el [13] 2013 for Groups C and D

load/ exp	X. Zhang et al[13]				Present Work			
	C-i	C-ii	D-i	D-ii	C-i	C-ii	D-i	D-ii
L-0	96.81	96.85	100	98.17	96.88	95.05	99.48	99.74
L-1	97.04	95.37	100	98.89	99.74	93.75	97.14	98.89
L-2	99.33	98.81	100	98.81	100	90.88	92.96	98.8
L-3	99.7	99.83	100	98.65	100	96.88	100	99.48

Group E is for a 10 class classification of all 3F of 3 FDs and N for each load condition. Although the authors in [2] have implemented a 10 class classification by LDA same as group E for a load of 3HP, using 9 time domain features and 5 time-frequency features for the vibration data of 48kHZ and have developed TR-LDA1 and TR-LDA2 algorithms to achieve 100% classification accuracy in classification and also has presented a comparison by implementing with different classifiers. It is observed LDA exhibits 98% of accuracy, and the same is achieved in present work using 4 features by LDA and NB classifiers. A similar implementation is performed by former authors for load conditions of 1 and 2 HP in [14] with the aid of K-means clustering and the implementations are identical to dataset E L-2 and E L-3 in present work. Correspondingly the authors have obtained 98.5% and 98.9% It is interesting to note that in the present work combination of four features are sufficient to obtain an AC of 92% as given in table 5, which once again illustrates that a combination of few good features are sufficient rather than a complicated algorithm for diagnosing the bearing faults, with minimum computational cost.

5. CONCLUSIONS

Fault diagnosis is a crucial part of condition monitoring of bearings to avoid unprepared repairs and cost-effective damages caused by failures. The condition monitoring scheme involves suitable feature extraction, feature reduction, feature selection and classification processes, among which feature extraction and classification play important role in the scheme. Once the feature extracted are effective enough to reveal all the characteristics of fault condition then feature reduction and feature selection processes can be evaded.

In this paper, for the first time statistical time domain features MAV, SSI, WL, WAMP, ZC, SSC is employed for the identification of the mechanical faults using LDA and NB classifiers. In this study the effectiveness of each feature is investigated pertaining to accuracy, sensitivity and specificity in 63 combinations for 15 datasets which are drawn from 5

groups. The FS of SSI, WL, SSC and ZC will contribute for maximum accuracy in all the cases. However, for many datasets FS of one, two and three features are giving better accuracy in which the features are WL, SSC and ZC. It is studied from the results that increasing the number of features will not contribute to improve the classification accuracy; instead if a feature represent time characteristics of a fault condition and another feature for frequency then their combinations are giving best accuracy results, conversely combining many features which are redundant in characteristics will not contribute much to improve classification accuracy therefore investigations are to be conducted to find the best feature combination and then employ classification with minimal number of features which reduces the overheads of dimensionality reduction schemes like feature selection and feature reduction. In addition, to this the present approach avoids the computational burden on classifiers. The low computational complexity of these features constitutes it a highly favorable feature to be employed as part of a system for real-time automated fault diagnosis schemes. The success of the present approach is verified through comparing the performance of classification problems from other researchers. It can be concluded that the features employed newly with NB classifier achieves more satisfactory results to discriminate the fault condition from vibration signal than the other methods do. While our system can achieve promising results for handling the fault diagnosis of roller bearings, our future work might focus on the following issues to improve the present approach (1) diagnosing the fault severities of each fault individually and even when present in combinations. (2) Diagnosing and investigating OR fault for different load zone conditions and to localize the OR faults. (3) Finally, to indicate the severity of fault and by defining certain fault level indicators, this aids in bearing performance prognostics in future.

REFERENCES

1. *Zhang P, Du Y, Habetler TG, Lu B (2011) A Survey of Condition Monitoring and Protection Methods for Medium-Voltage Induction Motors. IEEE Transactions on Industry Applications 47:34–46. doi: 10.1109/TIA.2010.2090839*
2. *Jin X, Zhao M, Chow TWS, Pecht M (2014) Motor Bearing Fault Diagnosis Using Trace Ratio Linear Discriminant Analysis. IEEE Transactions on Industrial Electronics 61:2441–2451. doi: 10.1109/TIE.2013.2273471*
3. *Prieto MD, Cirrincione G, Espinosa AG, et al (2013) Bearing Fault Detection by a Novel Condition-Monitoring Scheme Based on Statistical-Time Features and Neural Networks. IEEE Transactions on Industrial Electronics 60:3398–3407. doi: 10.1109/TIE.2012.2219838*
4. *Chebil J, Hrairi M, Abushikhah N (2011) Signal analysis of vibration measurements for condition monitoring of bearings. Australian Journal of Basic and Applied Sciences 5:70–78.*
5. *Bouchikhi E, Houssin E, Choqueuse V, et al (2012) Induction machine fault detection enhancement using a stator current high resolution spectrum. In: IECON 2012-38th Annual Conference on IEEE Industrial Electronics Society. IEEE, pp 3913–3918*
6. *Garcia-Perez A, Romero-Troncoso R de J, Cabal-Yepez E, Osornio-Rios RA (2011) The Application of High-Resolution Spectral Analysis for Identifying Multiple Combined Faults in Induction Motors. IEEE Transactions on Industrial Electronics 58:2002–2010. doi: 10.1109/TIE.2010.2051398*
7. *Blodt M, Chabert M, Regnier J, Faucher J (2006) Mechanical Load Fault Detection in Induction Motors by Stator Current Time-Frequency Analysis. IEEE Transactions on Industry Applications 42:1454–1463. doi: 10.1109/TIA.2006.882631*
8. *Peng ZK, Chu FL (2004) Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. Mechanical Systems and Signal Processing 18:199–221. doi: 10.1016/S0888-3270(03)00075-X*
9. *Jawadekar A, Paraskar S, Jadhav S, Dhole G (2014) Artificial neural network-based induction motor fault classifier using continuous wavelet transform. Systems Science & Control Engineering 2:684–690. doi: 10.1080/21642583.2014.956266*

10. Lou X, Loparo KA (2004) Bearing fault diagnosis based on wavelet transform and fuzzy inference. Mechanical Systems and Signal Processing 18:1077–1095. doi: 10.1016/S0888-3270(03)00077-3
11. Sugato Ghosh, Inertia Effect Under Couple Stress Fluid in Laminar Flow on Porous Journal Bearing, International Journal of Mechanical and Production Engineering Research and Development (IJMPERD), Volume 5, Issue 5, September - October 2015, pp. 1-12
12. Schmitt HL, Silva LRB, Scalassara PR, Goedtel A CLASSIFICATION OF SIMULATED MOTOR SIGNALS WITH BEARING FAULTS USING NEURAL NETWORKS, WAVELETS, AND PREDICTABILITY MEASURES.
13. Seshadrinath J, Singh B, Panigrahi BK (2014) Investigation of Vibration Signatures for Multiple Fault Diagnosis in Variable Frequency Drives Using Complex Wavelets. IEEE Transactions on Power Electronics 29:936–945. doi: 10.1109/TPEL.2013.2257869
14. Zhang X, Zhou J (2013) Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines. Mechanical Systems and Signal Processing 41:127–140. doi: 10.1016/j.ymssp.2013.07.006
15. Zhao M, Jin X, Zhang Z, Li B (2014) Fault diagnosis of rolling element bearings via discriminative subspace learning: Visualization and classification. Expert Systems with Applications 41:3391–3401. doi: 10.1016/j.eswa.2013.11.026
16. Zhang L, Xiong G, Liu H, et al (2010) Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference. Expert Systems with Applications 37:6077–6085. doi: 10.1016/j.eswa.2010.02.118
17. Samanta B, Al-Balushi KR, Al-Araimi SA (2003) Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. Engineering Applications of Artificial Intelligence 16:657–665. doi: 10.1016/j.engappai.2003.09.006
18. William PE, Hoffman MW (2011) Identification of bearing faults using time domain zero-crossings. Mechanical Systems and Signal Processing 25:3078–3088. doi: 10.1016/j.ymssp.2011.06.001
19. Wu S-D, Wu P-H, Wu C-W, et al (2012) Bearing Fault Diagnosis Based on Multiscale Permutation Entropy and Support Vector Machine. Entropy 14:1343–1356. doi: 10.3390/e14081343
20. Zhang S, Li W (2014) Bearing Condition Recognition and Degradation Assessment under Varying Running Conditions Using NPE and SOM. Mathematical Problems in Engineering 2014:1–10. doi: 10.1155/2014/781583
21. Raj AS, Murali N (2013) Early Classification of Bearing Faults Using Morphological Operators and Fuzzy Inference. IEEE Transactions on Industrial Electronics 60:567–574. doi: 10.1109/TIE.2012.2188259
22. Liu C, Wang G, Xie Q, Zhang Y (2014) Vibration Sensor-Based Bearing Fault Diagnosis Using Ellipsoid-ARTMAP and Differential Evolution Algorithms. Sensors 14:10598–10618. doi: 10.3390/s140610598
23. Amar M, Gondal I, Wilson C (2015) Vibration Spectrum Imaging: A Novel Bearing Fault Classification Approach. IEEE Transactions on Industrial Electronics 62:494–502. doi: 10.1109/TIE.2014.2327555
24. Smith WA, Randall RB (2015) Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. Mechanical Systems and Signal Processing 64-65:100–131. doi: 10.1016/j.ymssp.2015.04.021.
25. Case Western Reserve University Bearing Data Center Website. Available online: http://csegroups.case.edu/bearingdatacenter/pages/download-data-file/ (accessed on 23rd March 2015.